Application of Frailty Models on Advance Liver Disease Using Gamma as Frailty Distribution

Gurprit Grover, Divya Seth*

Department Of Statistics, Faculty of Mathematical Sciences, University Of Delhi, New Delhi - 110007, India

gurpritgrover@yahoo.com; divyaseth1234@gmail.com

Received 10 December, 2013; Revised 10 March, 2014; Accepted 20 March, 2014; Published 18 May, 2014

Abstract

A frailty model is a random effect model for time-to-event data, where the random effect has a multiplicative effect on the baseline hazard function. The notion of frailty provides a convenient way of introducing unobserved heterogeneity and associations into models for survival data. These are mixture models which are the extension of survival models considering two sources of variability: variability by observable risk factors included in the model and heterogeneity caused by unknown covariates. The purpose of introducing these models is to improve the fit of mortality models in population. This article aims to investigate the use of frailty model with concomitant variable in order to individualize the prognostic factors that influence the survival times of advance liver disease patients. A retrospective data of 811 admitted patients suffering from advance liver disease (cirrhosis and HCC) were analyzed with and without incorporating frailty effect to Cox PH and AFT model. It is conclusively established that as cites, weakness, edema, LGI bleed, abdominal distention and higher level of serum bilirubin and serum creatinine are found to be significant contributors under three considered models viz. No frailty, unshared frailty and shared frailty. Model selection was ascertained by Akaike's information criterion (AIC).Log logistic AFT model with shared frailty is found to be the best fit among all the three competing models.

Keywords

AFTM; AIC; Cirrhosis; Cox PH Model; Frailty Model

Introduction

Research in the field of survival analysis has increased greatly over the last several decades because of its large usage in areas related to biology, medicine, public health, and epidemiology. A typical analysis of survival data involves the modelling of time-to-event data, such as the time until death. Among the methods which estimate the survival function, several are widely used and their purpose is to compare the efficiency between competing estimators of the survival function. Although Cox proportional hazard (PH) model is widely applicable to analyze survival data, there are relatively few probability distributions for the survival time that can be used with these models. In these situations, the accelerated failure time (AFT) model is an alternative to the Cox PH model for the analysis of survival time data. Under AFT models we measure the direct effect of the explanatory variables on the survival time instead of hazard, as we do in the Cox PH model. Both Cox PH and AFT models describes the relationship between survival probabilities and a set of covariates but neither of them takes into account the unmeasured variability among subjects beyond that of measured covariates. If heterogeneity is suspected across the subjects beyond the measured covariates then the random term which can account for this variability is known as *frailty*.It originates from gerontology where it is used to indicate that frail people have an increased risk for morbidity and mortality (Duchateau and Janssen, 2008). Frailty models are one possible extension to the AFT model or other survival models that allow for such dependence.

Frailty models are potential choices for modeling unexplained heterogeneity in a population in time to event data. These models are essentially survival models with both fixed and random effect terms. The fixed effects comprise of the observed portion of the model whereas random effect term accounts for

unexplained variability in the model. In other words, the random component, or frailty, was initially designed to account for variability due to unobserved individual-level factors that is otherwise unaccounted for by the other predictors in the model but was expanded to describe heterogeneity among groups of individuals or within an individual. In this setting, controlling for unobserved heterogeneity was implemented with univariate frailty and shared frailty models, which estimate the baseline parameters, the coefficients of the prognostic factors and the variance of frailty (assumed to follow gamma distribution with mean 1 and variance σ^2 to be estimated) (Zarulli et al., 2013). Shared frailty models are the survival data analogue to random effects model. It is a model where frailties are no longer observation specific but instead are shared across groups of observations, thus causing those observations within the same group to be correlated (Cleves et al., 2008).The key idea of these models is that individuals have different frailties, and the most frail will die earlier than the less frail (Wienke A, 2011). This paper illustrates the application of frailty models to the advance liver disease dataset. A retrospective cohort of admitted patients suffering from advance liver disease viz. liver cirrhosis and HCC is taken into account to model the frailty effect. Cirrhosis and hepatocellular carcinoma (HCC) are the leading causes of death worldwide.

Cirrhosis is a condition in which the liver slowly deteriorates and malfunctions due to chronic injury. Scar tissue replaces healthy liver tissue, partially blocking the flow of blood through the liver. A healthy liver is able to regenerate most of its own cells when they become damaged. With end-stage cirrhosis, the liver can no longer effectively replace damaged cells. Cirrhosis is the 12thleading cause of death by disease, accounting for 27,000 deaths each year (Perz et al., 2006). Chronic infections with hepatitis B virus (HBV) and hepatitis C virus (HCV) are well-recognized risk factors for cirrhosis and liver cancer, but estimates of their contributions to worldwide disease burden have been lacking. Hepatocellular carcinoma/liver cancer (HCC) is one of the most common malignant tumours worldwide with an estimated 500000 to 1 million new cases per year (Wands and Blum, 1991). The incidence ranges from four cases per 100 000 populations in USA to 150 cases per 100 000 populations in parts of Africa and Asia where HCC is responsible for a large proportion of cancer deaths. A rise in the incidence of mortality from HCC has been observed in different countries(El-serag and Mason, 1999). Approximately 77% of deaths from HCC occur in developing countries. The prognosis of HCC is dismal with 5-year survival being 1%– 4%(Pisani et al., 1993).

This paper gives a new approach to know the presence of unobserved factors (heterogeneity) with the help of observed prognostic factors.Various studies have been conducted to study the effects of prognostic factors incorporating frailty effect in different diseases like kidney transplant, waiting time to first pregnancy, genetic trait etc. but its application on advance liver disease is still an unexplored avenue. Yue and Chan (1997) proposed a dynamic frailty model assuming gamma distributed frailty effect on animal carcinogenesis experiment and reported that it provides greater flexibility than the identical frailty model in accounting for serial dependence in the multivariate survival data. Manton et al.(Manton K, 1986) compared the inverse normal and gamma models, together with Gompertz and Weibull baseline hazards, in a study of survival at advanced ages, based on the data from US medicare insurance. Pan W (2001) adapted an EM-algorithm approach analogous to that in the frailty model for the Cox model and showed that through simulation its performance compares favorably with that of the marginal independence approach. Hangal DD (2009) discussed the use of different frailty distributions in a Weibull extended bivariate exponential regression model. Cui and Sun (2004) adopted a marginal survival model for multivariate survival data and imposed a shared univariate gamma frailty correlation structure to the correlated data. They developed a test for checking the adequacy of the gamma frailty distribution. Feng et al. (Feng S, 2005) applied multivariate frailty model to analyze dataset of deceased donor kidney transplant using Poisson variance structures. Donor age, cerebrovascular accident (CVA), renal insufficiency and history of hypertension have been identified as the risk factors associated with elevated risk of graft failure. The likelihood functions of both parametric and semi parametric multivariate frailty models are shown to be proportional to the likelihood functions of a class of mixed Poisson regression models. Chuang and Cai (2006) predicted the survival probabilities of patients with dental implant using Cox PH frailty model with clustering effect. Zareand Moradi (2012) applied parametric frailty and shared frailty models to waiting time to first pregnancy and found that height, age at marriage and menstruation regularity to be important predictors of waiting time to pregnancy (WTTP). Govindarajulu et al.(Govindarajulu US, 2011)

has applied the methodology to choose between frailty and no- frailty models in assessing genetic variability and found sex and birth year as significant covariates.

Grover et al. (Grover G, 2013) has proved that AFT model is better than Cox model in determining the predictors of the liver cirrhosis. However, this cannot be generalized to models where a random effect representing heterogeneity in the dataset is incorporated.

In this article, we intend to illustrate the performance of univariate and shared frailty model as an extension to Cox PH and AFT models. The aim of this study is firstly, to identify a frailty model that provides an adequate fit to advance liver disease dataset. Secondly, to enhance the understanding of the prognostic factors of advance liver disease by applying frailty models. Thirdly, to interpret the resulting parameter estimates in the context of the current epidemiologic and biological knowledge of advance liver disease. To the best of our knowledge, this is the first such investigation to apply frailty models for survival time of advance liver disease patients in the presence of predictors using Cox PH model and AFTM. Model choice criteria like log likelihood and AIC are used to choose the better performing model.

Data

For this study, patients admitted in Department of Gastroentrology of PushpawatiSinghania Research Institute, New Delhi, India who were suffering from liver cirrhosis and HCC were selected. A total of 811 patients were admitted for the treatment of liver cirrhosis and HCC from January 2007 to December 2010 and monitored through subsequent follow-ups. The institute is known for its expertise in liver and renal disease ailments and thus serves to a large segment of population from Northern region of India.

Materials and Methods

In the present study, we have considered three types of models viz. Model 1(No frailty), Model 2(Univariate/Unshared frailty) and Model 3(Shared frailty). Under model 1, Cox PH and AFT models are discussed without taking frailty into account whereas in model 2 and 3 survival models are analyzed using frailty effect. In model 1, we consider all 811 patients suffering from advance liver disease and analyzed them without using concept of frailty. In model 2, univariate frailty is applied to univariate survival data i.e. each cluster has only one individual with only one survival outcome. Kosork et al. (Kosork MR, 2004) addressed that univariate survival data can only have univariate frailty under certain conditions. Univariate frailty effect is assessed using two AFT models viz. Weibull and Loglogistic. In order to observe shared frailty effect (Model 3) in Cox PH and AFT models, we categorized patients enrolled in the study into two clusters namely, cirrhosis and HCC depending upon their initial condition. Under this model, univariate frailty is applied to multivariate survival data i.e. there is more than one individual in each cluster and all individuals share a common frailty within each cluster. The frailties across different clusters are assumed to have a distribution and account for unexplained heterogeneity at the cluster level. The frailty term, can be univariate or multivariate.

There are several methods to access the frailty effect. We choose gamma distribution to assess the frailty effect. According to Hougaard P(1995) the mathematical reasons to choose gamma as frailty distribution is that it has a simple density function for which parameters are easily obtained through likelihood estimation. He also discussed how the choice of frailty distribution was extended to the natural exponential family, where the gamma distribution is the simplest family of this kind. In gamma frailty models, the effect of covariate differences eventually diminishes completely in favor of frailty effect, i.e. at one point of time frailty becomes a major factor for the survival of the patient diminishing the presence of various prognostic factors (Duchateau and Janssen, 2008).

The predictors of interest are age, gender, haemoglobin (HB), total leukocyte count (TLC), platelet, serum creatinine, serum bilirubin, prothrombin time (PT), sodium, potassium, SGOT, SGPT, total protein, albumin, ascites, encephalopathy, fever, weakness, edema, anorexia, black stool, altered sensorium, portal hypertension, UGI bleed, LGI bleed and abdominal distention.

In general, the hazard function for the frailty model can be written as

$$h_{ij}(t \mid x, z) = h_0(t)c(\beta' x_{ij} + z'_{ij}\, w_{ij}) \qquad (1)$$

Where h_{ij} is the hazard function for the j^{th} individual from the i^{th} cluster, where cluster can be an individual or a group, $h_0(t)$ is the baseline hazard function and $\beta`$ is a vector of estimated coefficients associated with the observed covariates/prognostic factors vector x_{ij}, and w_{ij} is a vector of random effects associated with the covariates vector z_{ij} for cluster i and individual j, which

usually includes an intercept. The random effect term accounts for heterogeneity effects of the z_{ij}. When z_{ij} includes only an intercept the model (1) is a simple multiplicative frailty model. The w_{ij} term accounts for the correlation among individuals within a cluster.

Univariate/Unshared Frailty Model

Let us assume that there are n individuals (j=1, 2,….n_i) suffering from advance liver disease and their survival and hazard function are given by S(t) and h(t) respectively. And, let α represents the frailty effect (heterogeneity) which is an unobservable multiplicative effect on the hazard function assumed to follow some (gamma) distribution.

The individual hazard function for the j^{th} individual conditional on frailty effect αis given by

$$h_j(t \mid \alpha_j) = h_0(t) * \alpha_j \tag{2}$$

where $h_0(t)$ is the baseline hazard function and α_j denotes the frailty effect for the j^{th} individual.

The population/unconditional hazard function with gamma frailty is given by

$$h_\theta(t) = \frac{h(t)}{1-\theta \ln\{S(t)\}} \tag{3}$$

The individual survival function for the j^{th} individual conditional on frailty effect α is given by

$$S_j(t \mid \alpha_j) = \{S_j(t)\}^{\alpha_j} \tag{4}$$

whereα_j denotes the frailty effect of the j^{th} individual.

The population survival function of the advance liver disease patients is obtained by integrating out the unobservable α. If α has probability density function g(α), the population survival function is thus given by

$$S_\theta(t) = \int_0^\infty \{S(t)\}^\alpha g(\alpha) d\alpha \tag{5}$$

where θ represents the dependence of survival function on the frailty variance.

Since we have taken gamma distribution to explain the frailty effect, therefore if frailty effect α is distributed as gamma with mean 1 and variance θ, then

$$g(\alpha) = \frac{\alpha^{1/\theta - 1} e^{-\alpha/\theta}}{\Gamma(1/\theta)\theta^{1/\theta}} \tag{6}$$

And thus equation 5, the population survival function with gamma frailty can be written as

$$S_\theta(t) = [1-\theta \ln\{S(t)\}]^{-1/\theta} \tag{7}$$

We have considered two AFT models viz. Weibull and loglogistic for modellingunivariate frailty effect.

The larger the values of θ, greater is the heterogeneity between subgroups and a stronger association among members of a sub group is indicated. Also, individuals with a higher value of α will be more frail and have a high risk of developing the disease, whereas individuals with low value of α will be less frail and tend to survive longer with lower risk.

Shared Frailty Model

For the dataset consisting of n individuals, we divide them into two clustersviz. cirrhosis and HCC depending upon their initial condition. The individual hazard function of the j^{th} individual belonging to the i^{th} cluster generalizes to

$$h_{ij}(t \mid \alpha_i) = \alpha_i * h_{ij}(t) \tag{8}$$

where (j= 1,2….n_i: i = 1,2) with $h_{ij}(t) = h(t_{ij} \mid x_{ij})$. The frailties, α_i, are common (shared) within i^{th} cluster and are assumed to follow gamma distribution. The frailty varianceθ, is estimated from the data and measures the variability of the frailty across clusters.

In case of Weibullmodel, the conditional hazard for an individual is given by

$$h_{ij}(t \mid \alpha_i) = \alpha_i * \exp(x_{ij}\beta)\, p t^{p-1} \tag{9}$$

And the conditional survival function is given by

$$S_{ij}(t \mid \alpha_i) = \{S_{ij}(t)\}^{\alpha_i} = \exp\{-\alpha_i \exp(x_{ij}\beta) t^p\} \tag{10}$$

Similarly, for log logistic model the conditional hazard is given by

$$h_{ij}(t \mid \alpha_i) = \alpha_i * \frac{p t^{p-1}}{(1/\exp(x_{ij}\beta)) + t^p} \tag{11}$$

And conditional survival function is given by

$$S_{ij}(t \mid \alpha_i) = (1 + \exp(x_{ij}\beta) t^p)^{-1} \tag{12}$$

In univariate frailty models, the distribution for frailty is often given a prespecified fixed mean 1 for multiplicative hazards model to identify the frailty distribution to overcome the issue of identifiability. Identifiability refers to being able to uniquely estimate both the parameters of the hazard function as well as of the frailty distribution in univariate data (Pickles and Crouchley, 1995).

Estimation of Parameters

There are various methods to estimate the parameters of the frailty models namely, Expectation Maximization (EM) algorithm, penalized likelihood, Newton-Raphson, MCMC (Monte Carlo Markov Chain) etc. Depending upon the particular distribution specified for the frailty random effect, the frailty is

integrated out of the joint likelihood to obtain an observed likelihood as afunction of survival time given the covariates. After obtaining a likelihood with the frailty terms integrated out, a numerical procedure like Newton-Raphson is used for maximum likelihood parameter estimation.

The likelihood of the observed data is obtained by calculating the group- level conditional likelihoods and integrating out frailty. Suppose we have data for i=1,2 and j= 1,2 ….n_i, observations per clusterconsisting of trivariate response(t_{0ij}, t_{ij}, d_{ij}), which indicate the start time, end time and failure time for the j^{th} individual for the i^{th} cluster. If we define $D_i = \sum_{j=1}^{n_i} d_{ij}$ then, the contribution of the j^{th} individual to the conditional likelihood L_i, when the frailties follow a gamma distribution, L_i can be compactly expressed as

$$L_i = \left[\prod_{j=1}^{n_i}\{h_{ij}(t_{ij})\}^{d_{ij}}\right] \frac{\Gamma(1/\theta + D_i)\theta^{D_i}}{\Gamma(1/\theta)}\left[1-\theta\sum_{j=1}^{n_i}\ln\frac{S_{ij}(t_{ij})}{S_{ij}(t_{0ij})}\right]^{-1/\theta-Di} \quad (13)$$

Given the unconditional group likelihoods, we can estimate the regression parameters and frailty variance θ by maximizing the overall log-likelihood

$$\ln L = \sum_{i=i}^{n} \ln L_i \quad (14)$$

The details of estimation and inference procedures of the models are well explained by Gutierrez (Gutierrez, 2002). Model selection was ascertained by comparing log likelihood values and on the basis of Akaike's information criterion(AIC) (Akaike, 1974). Lower AIC value gives the best model fit.

Choosing Between Frailty and Non-frailty Models

One can incorporate a frailty term into a parametric or semi parametric model where the model can either be univariate or bivariate or multivariate/ correlated. If the frailty term is not significant in the model, it implies that there is no significant frailty effect and, therefore, a frailty model for the data is not considered to be distinctly different from a non-frailty model. In case the frailty term is significant in the model, one could proceed further using model validation to compare how well the frailty model compares with a non-frailty model.

Results

A total number of 811 admitted patients suffering from Cirrhosis and HCC were studied retrospectively. Out of which, 657 (81%) and 154 (19%) were identified as patients suffering from cirrhosis and HCC respectively. The details of descriptive statistics have been discussed in Grover et al. [18]. 13 patients with co-morbidities like renal failure, cardiac arrest, gastric cancer etc are excluded from the study because these numbers were not sufficient to draw valid conclusions about these identities.

In this study, in addition to the Cox proportional hazard model, Weibull AFTM and Loglogistic AFTM are considered with frailty effects for analyzing survival data and for examining the effects of various prognostic factors on the survival of patients with advance liver disease. The frailty models are considered for adjusting the heterogeneity in advance liver disease at the patient level. In all considered models, the frailty effect was observed for those variables that were statistically significant at 0.10 level in univariate analysis whereas during multivariate analysis 5% was the considered level of significance. We implemented the models in statistical software STATA 11.0. Modeling a frailty effect is not only a function of unobserved heterogeneity but also of observed covariates. So, we present the result of observed covariates which is a novel application of frailty models to address the limitations of existing survival models.

Table 1 shows the estimated hazard ratio for Cox PH model and estimated time ratio, corresponding to the parameters and the corresponding coefficients for all the three models considered in this analysis. Cox PH model, weibull AFTM and loglogistic AFTM are compared without frailty and with shared frailty effect whereas only AFT models are compared under univariate frailty effect.

In model 1, the fit of Cox PH model and other AFT models without frailty effect shows that almost all the covariates considered in the model have significant effect on the likelihood of advance liver disease. It can be observed that after the onset of advance liver disease the males face the hazard 67% higher than females. Also, ascites, abdominal distension, weakness and serum Bilirubin are the significant contributors in the mortality of the patients initially diagnosed with advance liver disease. Patients with symptoms ofascites, abdominal distention, weakness, edema, UGI bleed and higher level of serum bilirubin are likely to face 93%, 24%, 83%, 21%, 5% and 3% greater hazard than the patients without these symptoms. Also, male gender, ascites, abdominal distention, weakness, and

serum bilirubin are the significant contributors in the survival of the patients.

In model 2, considering unshared frailty effect comparison of Weibull AFT model and loglogistic AFT model are carried out. We noted that under this model ascites, weakness, edema, and higher level of serum Bilirubin are the significant contributors. Thus, the estimated deceleration factor suggests that the presence of ascites, weakness, edema and higher level of serum Bilirubin are significantly effective for hastening failure/decreasing survival time by a factor of 0.34, 0.54, 0.66 and 0.97 respectively. A significant frailty effect is observed in both the models under univariate without shared frailty. Thus, if we believe the individual hazard to be Weibull or loglogistic then we must also be willing to believe in an unobserved patient level effect, the value of $\theta \neq 0$ and is highly significant. So, there is a sufficient evidence, to believe that heterogeneity is observed in the community at patient level itself. On comparing AIC of both the AFT models; loglogistic model as a better fit in a univariate without shared frailty model.

Since it is possible to expect some correlation within a cluster (i.e cirrhosis and HCC), we have depicted this in model 3, as shared frailty model where sharing takes place on the cluster level. The shared function is fit to the status of the patients either cirrhosis or HCC, which is used as the cluster number for fitting a shared frailty model. Few additional significant prognostic factors are observed to be significant while carrying out the analysis under shared frailty. In addition to, ascites and weakness other prognostic factors like LGI bleed and altered sensorium are also found to be significant contributors in the survival of the patients while controlling for cluster level effect.

To test the significance of frailty effect, we evaluated the value of theta which is quiet high as shown in Table1. The test statistic is statistically significant at a 0.05 significance level, thus there is sufficient evidence to reject H_0, which in essence says that shared frailty model serves our purpose well. According to AIC criterion, loglogistic model is found to be the best fit among all other competing models under shared frailty effect.

TABLE 1.COMPARISON OF THE COEFFICIENTS AND THEIR CORRESPONDING HAZARD/TIME RATIOS OF VARIOUS SURVIVAL MODELS WITH AND WITHOUT FRAILTY EFFECT.

Prognostic factors	Model 1 (Without Frailty)						Model 2 (Unshared Frailty)				Model 3 (Shared Frailty)					
	Cox		Weibull		Loglogistic		Weibull		Loglogistic		Cox		Weibull		Loglogistic	
	Coeff	HR	Coeff	TR	Coeff	TR	Coeff	TR	Coeff	TR	Coeff	HR	Coeff	TR	Coeff	TR
Sex	0.513*	1.67	-0.595*	0.55	-0.625*	0.53	-	-	-	-	0.425	1.53	-0.407	0.66	-0.446	0.64
Ascites	0.657*	1.93	-0.756*	0.47	-0.893*	0.41	-1.064*	0.34	-1.037*	0.35	0.673*	1.96	-0.635*	0.53	-0.864*	0.42
Abdominal distention	0.215	1.24	-0.261	0.77	-0.246	0.78	-0.201	0.82	-0.196	0.82	-	-	-	-	-	-
Weakness	0.604*	1.83	-0.717*	0.49	-0.720*	0.49	-0.618*	0.54	-0.579*	0.56	0.351*	1.42	-0.354*	0.70	-0.313	0.73
Edema	0.191	1.21	-0.228	0.80	-0.272	0.76	-0.419*	0.66	-0.374	0.69	-	-	-0.263	0.77	-0.446*	0.64
UGI Bleed	0.048	1.05	-0.072	0.93	-0.005	0.99	-0.062	0.94	-0.174	0.84	-	-	-	-	-	-
Serum Bilirubin	0.034*	1.03	-0.030*	0.97	-0.026*	0.97	-0.030*	0.97	-0.031*	1.03	0.019	1.02	-0.020	0.98	-0.021*	0.97
SGPT	-	-	-	-	-	-	-0.003	0.99	-0.003	0.99	-	-	-	-	-	-
LGI Bleed	-	-	-	-	-	-	-	-	-	-	0.601*	1.82	-0.615*	0.54	-0.463	0.63
Serum Creatinine	-	-	-	-	-	-	-	-	-	-	0.131*	1.14	-0.143*	0.87	-0.156*	0.86
SGOT	-	-	-	-	-	-	-	-	-	-	-0.010	0.99	0.007	1.01	0.001	1.01
Altered Sensorium	-	-	-	-	-	-	-	-	-	-	0.626*	1.87	-0.638*	0.53	-0.789*	0.45
Theta (θ)	-	-	-	-	-	-	5.191#	-	2.078#	-	0.6627#	-	0.6842#	-	0.7693#	-
-2Log likelihood	2321.64		1232.98		1222.94		1179.8		1178.74		2102.78		1066.04		1051.38	
AIC	2328.64		1242.98		1231.94		1188.8		1187.74		2109.78		1076.04		1061.38	

*Significant at 5%

Significant at 1%

Discussion

The structure, properties, and applicability of survival models to practical problems depend on the nature of time-to-event data, ancillary information about influential factors and the aims of the studies. Ordinary survival models deal with the simplest case of independent and identically distributed data. This is

based on the assumption that the study population is homogeneous. But it is a basic observation of medical statistics that individuals differ greatly, so do the effects of a drug, or the influence of various explanatory variables. This heterogeneity is generally recognized as one of the most important sources of variability in medical and biological applications. In a heterogeneous population, the population hazard can fall while the individual hazard may rise because, over time, the population becomes populated by more and more robust individuals as the frailer member fail. It virtually ensures that population hazards decline over time, regardless of the shape of the hazards that individual face (Cleves et al., 2008, Chapter 15). With the passage of time, the covariates become less of a factor and frailty becomes more of a factor in determining the chance of remaining infection/disease free.

These types of model are used particularly in medical applications in which stages or levels are recurrent in nature, correlated data, biomedical data etc. Such models have been used in a wide range of medical applications, for instance kidney transplant (Feng, 2005), birth interval (Mahmood, 2013), Hodgkin lymphoma (Grotmol, 2011), biomedical and genetic studies (Govindarajulu, 2011),breast cancer (Raman and Venkatsan, 2012) etc. but in a very limited range to non medical application like mortality by education (Zarulli, 2013). All these reported studies have applied univariate, bivariate or multivariate/correlated frailty models to univariate or multivariate survival data accordingly.

Analysis of multivariate survival data provides an exciting example for challenging modeling strategies. To illustrate the strength of frailty models, we have proposed a frailty-based approach for estimation in the linear regression (i.e Cox PH and AFT model) for multivariate failure time data of advance liver disease which explicitly takes into account the possible correlation among failures times. Various prognostic factors which are affecting the survival of the advance liver disease patients have been identified and interpreted in the terms of current epidemiologic and biological knowledge. Many studies have evaluated the risk factors for advance liver disease using non-parametric, semi parametric or parametric approach but none of them discussed ahead have used frailty based approach. Livraghi et al. (Livraghi, 1995) reported 3 – 5 year survival rate of cirrhosis an HCC patients who had undergone PEI (Percutaneous Ethanol Injection) using Kaplan Meier technique and found that patients with Child A, B, C cirrhosis and single HCCs of 5 cm or smaller, their survival rate was 47%-79% and 0%-12% respectively.

The results of frailty models have confirmed much of what is known about the natural course and the factors affecting liver cirrhosis and HCC or combining them to call advance liver disease. However, using frailty models we have learned more about how the different prognostic factors explains unobserved heterogeneity which differs at individual level and cluster level. We observed the predictors namely, sex, ascites, weakness, edema, serum Bilirubin, abdominal distention, LGI bleed and serum creatininehave significant impact on the survival time of the patients suffering from advance liver disease. Our findings coincide with the reports of Samada et al. (Samada M, 2008), in which it has been reported that male sex, presence of ascites and higher level of serum bilirubin are associated with an increased risk of death. Male subjects are likely to face the hazard 53% more as compared to female subjects, while presence of ascites and weakness decreases the survival time of patients by 42% and 73% respectively. Reporting of weakness is subjective in nature and may vary from individual to individual but from author point of view it can be a contributing cause for heterogeneity.

Serum Bilirubin is another significant factor contributing to poor survival. Shapiro et al.(Shapiro JM, 1979) reported that whenever two successive values of serum Bilirubin taken six months apart exceeded 2 mg/dl then patient had entered a late phase of liver disease and on an average survived for 49 months.

Attia et al. (Attia K A, 2008) reported creatinine as an independent predictor of mortality in African patients suffering from cirrhosis. Our results incorporating frailty effect confirmed the author's findings and indicated lower survival of patients.

Similar to our findings, Nouso et al. (Nouso K, 2008) observed serum Bilirubin, presence of uncontrollable asites, and a high platelet count as well as multiple tumours, large tumours(>3cm), high alpha-fetoprotein and presence of portal vein thrombosis as the factors for poor prognosis of HCC(Hepatocellular carcinoma).

The hazard ratio for men versus women deceases by more than 10% when frailties are included in the model as compared to model with no frailties. Similar pattern is observed with covariate weakness, hazard ratio for patients with weakness versus without weakness decreases by 40% when frailties are included

in the model as compared to model with no frailties. As far as, ascites and serum Bilirubin is concerned no such difference is observed.

The variance components are modeled on the log-hazard scale. Pankratz et al. (Pankratz, 2005) described that exponentiating the square root of variance component provides information concerning the relative of the outcome that corresponds to the random effect. The estimated variances of frailty effect in model 2 are 5.19 and 2.08 for weibull and loglogistic respectively. Thus, exp (√5.19) = 9.75, so most individual relative risk of dying at a given point of time are up to 9.75 time larger or smaller than the overall risk. Also, the estimated variances for the frailty effects in model 3are θ = 0.6627, 0.6842, 0.7693 for Cox PH and AFT models under shared frailty respectively. So, exp(√0.6627) = 2.26, most individual cluster will have risk of death at a given point of time are up to 2.26 times larger or smaller than the overall risk. The comparison between univariate frailty and shared frailty is based the comparison of log likelihoods keeping in mind that the shared frailty model contains one parameter less (Wienke A, 2011, Chapter 4).

The main message from this study is that properly taking into account the correlation among failure times can substantially improve the efficiency in estimating regression coefficients, especially when the within cluster correlation is strong.Frailty modelling can contribute to a clarification in the disagreement apparently existing between those suggesting that all individuals are at risk to develop liver cancer, but with heterogeneity in the risk. The results of this frailty modelling indicate that it is possible to model the incidence of liver cancer in the population based on large heterogeneity in risk between individuals, in such a manner that a small group of individuals are susceptible to develop the disease, whereas the remaining majority have a low risk.

REFERENCES

Akaike, H. "A new look at the statistical model identification." IEEE Transactions on Automatic Control 19 (1974): 716-723.

Attia, K. A., Ackoundou-N`guessan, K. C., N`dri-Yoman, A. T., Mahassadi, A. K., Messou, E, Bathaix, Y. F. and Kissi, Y.H."Child-Pugh-Turcott versus Meld score for predicting survival in a retrospective cohort of black African cirrhotic patients." World Journal of Gastroenterology 14(2) (2008): 286-291.

Chuang, S. K., Cai, T. "Predicting clustered dental implant survival using frailty methods." Journal of dental research 85(12) (2006): 1147-1151.

Cleves, M. A., Gould, W. W., Gutierrez, R. G. and Marchenko, Y. U. An introduction to survival analysis using STATA.2nd edition, Stata press publication, Texas, 2008.

Cui, S. and Sun, Y. "Checking for the gamma frailty distribution under the marginal proportional hazards frailty model."Statistica Sinica 2004, 14 (2004): 249-267.

Duchateau, L. and Janssen, P. The frailty Model, Springer, New York, 2008.

El-serag, H. B. and Mason, A.C."Rising incidence of hepatocellular carcinoma in the United States." New England Journal of Medicine 340 (1999): 745–750.

Feng, S., Wolfe, R. A. and Port, F. K. "Frailty survival model analysis of the national deceased donor kidney transplant dataset using Poisson variance structures." Journal of American Statistical Association100 (471) (2005): 728-735.

Govindarajulu, U. S., Lin, H., Lunetta, K. L. and Agostini, RBD Sr. "Frailty Models: Applications to biomedical and genetic studies." Statistics in Medicine 30 92011): 2754-2764.

Grotmol, T., Bray, F., Holte, H., Haugen, M., Kunz, L., Tretli, Aalen, O. O. and Moger, T. A. "Frailty modeling of the bimodal age- incidence of Hodgkin lymphoma in the Nordic countries." Cancer epidemiology, biomarkers & prevention 20(7) (2011): 1350-1357.

Grover, G., Sreenivas, V., Khanna, S. and Seth, D. (2013). "Estimation of survival of liver cirrhosis patients in the presence of prognostic factors using accelerated failure time model as an alternative to proportional hazard model." International journal of statistics and applications 3(4) (2013): 113-122.

Hangal, D. D. "Weibull extension of bivariate exponential regression model with different frailty distributions." Statistical Papers 50(1) (2009): 29-49.

Hougaard, P. "Frailty models for survival data." Lifetime data analysis 1 (2009): 255-273.

Kosorok, M. R., Lee, B-L.and Fine, J.P. "Robust inference for univariate proportional hazards frailty regression models." The Annals of Statistics 32(4) (2004): 1448-1491.

Livraghi, T., Giorgio, A., Marin, G., Salmi, A., de Sio,

I., Bolondi, L., Pompili, M., Brunello, F., Lazzaroni, S., Torzilli, G. "Hepatocellular carcinoma and cirrhosis in 746 patients: long-term results of percutaneous ethanol injection". Radiology, 197(1) (1995): 101-108.

Mahmood, Sharif., Zainab, B. and Mahbub, Latif. A.H.M. (2013). Frailty modeling for clustered survival data: an application to birth interval in Bangladesh. Journal of applied satatistics http://dx.doi.org/10.1080/02664763. 2013.825702.

Manton, K., Stallard, E. and Vaupel, J. "Alternative models for heterogeneity of mortality risks among the aged." Journal of the American Statistical Association 81 (1986): 635-644.

Nouso, K., Ito, Y. M., Kuwaki, K., Kobayashi, Y., Nakamura, S., Ohashi, Y. and Yamamoto, K. "Prognostic factors and treatment effects for hepatocellular carcinoma in Child C cirrhosis." British Journal of Cancer 98 (2008): 1161-1165.

Pan, W. "Using frailties in the accelerated failure time model." Lifetime data analysis 7 (2001): 55-64.

Pankratz, V. S., de Andrade, M. and Therneau, T. M."Random-effects Cox proportional hazard model: general variance components methods for time-to-event data." Genetic Epidemiology, 28 (2005), 97-109.

Perz, J. F., Armstrong, G. L., Farrington, L. A., Hutin, Y. J. F and Bell, B.P. "The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide." Journal of Hepatology 45(4) 92006): 529-538.

Pickles, A. and Crouchley, R. "A comparison of frailty models for multivariate survival data." Statistics in Medicine 14 (1995): 1447-1461.

Pisani, P., Parkin, D. M. and Ferlay, J."Estimates of the worldwide mortality from eighteen major cancers in 1985: Implications for prevention and projections of future burden." International Journal of Cancer 55 (1993): 891–893.

Raman, T. T. and Venkatsan, P. "Accelerated failure time frailty model in survival analysis." International journal of science and technology 2(2) (2012): 65-69.

Samada, M., Hernandez, J. C., Ramos, L., Barroso, L., Gonzalez, L., Cepero, Valdes. M., Hernandez, R. H., Abdo, C. A., Valdes, A. R., Perez, B. J. and Bernardos, A. "Factors that predict survival in patients with cirrhosis considered for liver Transplantation." Transplantation Proceedings 40 (9) (2008): 2965-2967.

Shapiro, J. M., Smith, H. and Schaffner, F. "Serum Bilirubin: a prognostic factor in primary biliary cirrhosis." GUT 20(2) (1979): 137-140.

Tito, L., Antonio, G., Giuseppe, M., Andrea, S., Ilario de, S., Luigi, B., Maurizio, P., Franco, B., Sergio, L., Guido, T. and Alberto Zucchi."Hepatocellular Carcinoma and Cirrhosis in 746 patients: Long-term Results of Percutaneous Ethanol Injection." Radiology197 (1995): 101-108.

Wands, J. R. and Blum, H. E."Primary hepatocellular carcinoma." New England Journal of Medicine 325 (1991):729–731.

Wienke, A. Frailty models in Survival Ananlysis, Chapman & Hall/CRC Biostatistics Series, 2011.

Yue, H. and Chan, K. S."A dynamic frailty model for multivariate survival data." Biometrics53 (1997): 785-793.

Zarulli, V., Marinacci, C., Costa, G. and Caselli, G. (2013). Mortality by education level at late-adult ages in Turin: a survival analysis using frailty model with period and cohort approaches. BMJ Open; 3:e002841. doi:10.1136/bmjopen-2013-002841.

Zare, N. and Moradi, F. "Parametric frailty and shared frailty models applied to waiting time to first pregnancy." International conference on applied mathematics and pharmaceutical sciences, 2012: 598-600.